Large deviation rates for branching processes

Li Junping
Joint work with Cheng Lan, Chen Anyue and Li Liuyan

Central South University

The 15th Workshop on Markov Processes and Related Topics
July, 11-15, 2019, Jilin Uni. Changchun

Contents

Contents

Contents

Contents

Contents

Background

- Branching process

State space $\mathbb{Z}_+ = \{0, 1, \cdots\}$.

▶ Definition

A conservative Q-matrix $Q = (q_{ij}; i, j \in \mathbb{Z}_+)$ is called a branching Q-matrix if it takes the following form:

$$q_{ij} == \begin{cases} ib_{j-i+1}, & \text{if } i \geq 1, j \geq i-1 \\ 0, & \text{otherwise,} \end{cases} \tag{1.1}$$

where

$$b_j \geq 0 \ (j \neq 1), \ 0 < -b_1 = \sum_{j \neq 1} b_j < \infty. \tag{1.2}$$

Background

A Branching Markov process (simply, MBP) is a continuous-time Markov chain taking values in $\mathbb{Z}_+$ whose transition function $P(t) = (p_{ij}(t); i, j \in \mathbb{Z}_+)$ satisfies the Kolmogorov equations

$$P'(t) = P(t)Q, \tag{1.3}$$

where Q is a branching Q-matrix.

Background

Let $\{X(t); t \geq 0\}$ denote the corresponding process and $P(t) = (p_{ij}(t); i, j \in \mathbf{Z}_+)$ denote its transition function. Define

$$F(t, u) = \sum_{j=0}^{\infty} p_{1j}(t) u^j.$$

It is well known that

$$\sum_{k=0}^{\infty} p_{ik}(t) u^k = [F(t, u)]^i.$$

Background

We assume throughout this talk that

$$b_0 = 0, \quad m =: \sum_{j=1}^{\infty} j b_{j+1} < \infty.$$

Therefore, Q is regular. It is also easy to see that

$$X(t) \to \infty \quad as \quad t \to \infty.$$

By Athreya and Ney [3] or Harris [2] that $W(t) = e^{-mt}X(t)$ is an integrable martingale and thus converges to a $r.v.$ W w.p.1 as $t \to \infty$. Hence, $\frac{X(t+s)}{X(t)}$ converges to e^{ms} w.p.1 as $t \to \infty$ for any fixed $s \geq 0$.

Background

- Problems: What are the following convergence rates?

$$P(|\frac{X(t+s)}{X(t)} - e^{ms}| > \varepsilon), \qquad (1.4)$$

and

$$P(|W(t) - W| > \varepsilon), \quad P(|\frac{X(t+s)}{X(t)} - e^{ms}| > \varepsilon | W \geq \alpha) \qquad (1.5)$$

for $\varepsilon > 0$ and $\alpha > 0$.

Background

- Related conclusions:

(i) G-W process: Athreya, K.B. (1994, Annals of Applied Probability, 4(3):779 – 790)

(ii) G-W process with immigration: (1.4), Liu J.N. and Zhang M. (2016, Acta Mathematica Sinica, English Series, 32(8):893-900).

(iii) G-W process with immigration: (1.5), Li L.Y. and Li J.P. (2018, Submitted to JOTP).

Preliminary

Define

$$B(u) = \sum_{j=0}^{\infty} b_j u^j.$$

Denote

$$g_k(t) = e^{-b_1 t} p_{1k}(t),\ k \geq 1, \quad H_{j,k}(t) = e^{j b_1 t} \int_0^t g_k(s) e^{-j b_1 s} ds,\ \ j \geq k.$$

Lemma 2.1.

For any $k \geq 1$,

$$g_k'(t) \geq 0, \qquad H_{j,k}'(t) \geq 0, \quad j \geq k.$$

Preliminary

Lemma 2.2.

For any $j \geq 1$,

$$\lim_{t\to\infty} e^{-b_1 t} p_{1j}(t) = \rho_j$$

exists and $\rho_j \leq \rho_1 = 1$ $(j \geq 1)$. Furthermore, $Q(u) = \sum_{j=1}^{\infty} \rho_j u^j$ is the unique solution of

$$B(u)Q'(u) - b_1 Q(u) = 0, \quad 0 \leq u < 1 \qquad (2.1)$$

subject to

$$Q(0) = 0, \; Q'(0) = 1, \; Q(u) < \infty \; (u \in (0,1), \; Q(1) = \infty. \qquad (2.2)$$

Idea of proof. Use Lemma 2.1 and Kolmogorov forward equation.

Preliminary

Note that $F(s,u)$ is strictly increasing with respect to $u \geq 0$.
For any fixed $s > 0$, let $g(s,\cdot) = F^{-1}(s,\cdot)$ be the inverse of $F(s,\cdot)$.
Denote

$$u_s = \sup\{u \geq 0;\ F(s,u) < \infty\},$$

then $u_s \geq 1$ and $g(s,\cdot)$ is well defined on $[0, F(s,u_s))$ (or $[0, F(s,u_s)]$ if $F(s,u_s) < \infty$) with value in $[0, u_s)$ (or $[0, u_s]$ if $F(s,u_s) < \infty$).

Preliminary

Lemma 2.3.

(i) $u_{s+t} \leq u_s$ and $F(s+t, u_{s+t}) \geq F(s, u_s)$ for any $s, t > 0$.
(ii) For any $s, t \geq 0$, $u \in [1, F(s, u_s))$,

$$g(s+t, u) \leq g(s, u).$$

(iii) For any $s, t \geq 0$, $u \in [0, 1]$,

$$g(s+t, u) \geq g(s, u).$$

Furthermore,

$$g(s+t, u) = g(t, g(s, u)) \quad u \in [0, F(s, u_s)).$$

Preliminary

Proposition 2.1.

If $F(s_0, u_0) < \infty$ for some $u_0 > 1$ and $s_0 > 0$, then for $1 \leq u \leq F(s_0, u_0)$, $g(t, u) \downarrow 1$ as $t \uparrow \infty$ and

$$R(t, u) \equiv e^{mt}(g(t, u) - 1) \downarrow R(u) \quad as \quad t \uparrow \infty \tag{2.3}$$

where $R(\cdot)$ is the unique solution of the functional equation

$$R(F(s_0, u)) = e^{ms_0} R(u) \quad for \quad 1 \leq u \leq F(s_0, u_0) \tag{2.4}$$

subject to

$$\begin{aligned} &0 < R(u) < \infty \quad for \quad 1 < u \leq F(s_0, u_0) \\ &R(1) = 0, \quad R'(1) = 1. \end{aligned} \tag{2.5}$$

Idea of proof. Use Lemma 2.3 and the properties of $F(s, u)$.

Conclusions

Theorem 3.1.

Suppose that $B(\theta_0) < \infty$ for some $\theta_0 > 1$. Then for any $\varepsilon > 0$ and $s > 0$,

$$\lim_{t\to\infty} e^{-b_1 t} P(|\frac{X(t+s)}{X(t)} - e^{ms}| > \varepsilon | X(0) = 1) = \sum_{k=1}^{\infty} \phi(s,k,\varepsilon)\rho_k < \infty \quad (3.1)$$

where $\phi(s,k,\varepsilon) = P(|\bar{Z}_k(s) - e^{ms}| > \varepsilon)$ and $\bar{Z}_k(s)$ being the mean $\frac{\sum_{i=1}^{k} Z_i(s)}{k}$ of k i.i.d. r.v. $Z_i(s)$ with same distribution as $X(s)$. $\{\rho_k\}$ is given in Lemma 2.2.

Conclusions

Sketch of proof. (i) By the condition, we can prove that for any $s > 0$, there exists $\tilde{\theta}_0 \in (1, \theta_0)$ $s.t.$ $F(s, \tilde{\theta}_0) < \infty$.
(ii) $X(t+s)$ can be expressed as

$$X(t+s) = \sum_{k=1}^{X(t)} \xi_{t,i}(s) \tag{3.2}$$

where $\{\xi_{t,i}(s);\ t \geq 0;\ i \geq 1\}$ are i.i.d. processes with the same law as $X(s)$. Therefore,

$$\begin{aligned} & P(|\frac{X(t+s)}{X(t)} - e^{ms}| > \varepsilon | X(0) = 1) \\ = & \sum_{k=1}^{\infty} P(X(t) = k | X(0) = 1) \phi(s, k, \varepsilon). \end{aligned}$$

Conclusions

(iii) Estimate Since $\phi(s,k,\varepsilon)$.
For any fixed $s>0$,

$$\begin{aligned} & \phi(s,k,\varepsilon) \\ \leq\ & P(\alpha^{\sum_{i=1}^k Z_i(s)} > \alpha^{k(e^{ms}+\varepsilon)}) + P(\beta^{\sum_{i=1}^k Z_i(s)} > \beta^{k(e^{ms}-\varepsilon)}) \\ \leq\ & [F(s,\alpha)\alpha^{-(e^{ms}+\varepsilon)}]^k + [F(s,\beta)\beta^{-(e^{ms}-\varepsilon)}]^k \end{aligned}$$

where α and β are any constants in $(1,\tilde{\theta}_0)$ and $(0,1)$ respectively.
(iv) It can be verified that for any $\varepsilon\in(0,1)$ and $s>0$, there exist $\alpha_0\in(1,\tilde{\theta}_0)$ and $\beta_0\in(0,1)$ $s.t.$

$$0 < F(s,\alpha_0)\alpha_0^{-(e^{ms}+\varepsilon)} < 1 \ and \ \ 0 < F(s,\beta_0)\beta_0^{-(e^{ms}-\varepsilon)} < 1.$$

Therefore, there exists $\lambda=\lambda(s,\varepsilon)\in(0,1)$ $s.t.$

$$\phi(s,k,\varepsilon) \leq 2\lambda^k, \quad \forall k\geq 1.$$

By Lemma 2.2, we get the result.

Conclusions

Remark.

Suppose that for fixed $s > 0$ and $\varepsilon > 0$, there exist constants $C_\varepsilon(s)$ and $r > 0$ such that $mr > -b_1$ and $\phi(s,k,\varepsilon) \le C_\varepsilon(s) \cdot k^{-r}$ for all $k \ge 1$. Then (3.1) holds.

Sketch of proof. (i) Note that

$$h(t,k) \equiv \frac{\phi(s,k,\varepsilon)P(X(t)=k)}{e^{b_1 t}} \le \frac{C_\varepsilon(s)}{k^r} \cdot \frac{P(X(t)=k)}{e^{b_1 t}} \equiv \tilde{h}(t,k).$$

By Lemma 2.2,

$$\lim_{t\to\infty} h(t,k) = \phi(s,k,\varepsilon)\rho_k$$

and

$$\lim_{t\to\infty} \tilde{h}(t,k) = C_\varepsilon(s) \cdot \frac{\rho_k}{k^r}.$$

Conclusions

(ii) Estimate $\sum_{k=1}^{\infty} \tilde{h}(t,k)$.

$$\begin{aligned}\sum_{k=1}^{\infty} \tilde{h}(t,k) &= \frac{E[X^{-r}(t)]}{e^{b_1 t}} \\ &= \frac{1}{\Gamma(r)} \int_0^{\infty} \frac{F(t,e^{-v})}{e^{b_1 t}} v^{r-1} dv \\ &= \frac{1}{\Gamma(r)} \int_0^{1} \frac{F(t,u)}{e^{b_1 t}} k(u) du \uparrow \int_0^1 Q(u)k(u)du.\end{aligned}$$

where $k(u) = \frac{|\log u|^{r-1}}{u}$. We have used $\frac{F(t,u)}{e^{b_1 t}} \uparrow Q(u)$ as $t \uparrow \infty$ which is due to Lemmas 2.1 and 2.2.

Conclusions

(iii) Prove that $\int_0^1 Q(u)k(u)du < \infty$.
Fixed $0 < u_0 < 1$, denote $u_n = g(n, u_0)$, it can be proved that

$$\int_{u_0}^{1} Q(u)k(u)du = \sum_{n=1}^{\infty} \int_{u_{n-1}}^{u_n} Q(u)k(u)du$$

and there exists $\lambda \in (e^{-(mr+b_1)}, 1)$ $s.t.$ for n large enough,

$$\int_{u_n}^{u_{n+1}} Q(u)k(u)du \leq \lambda \int_{u_{n-1}}^{u_n} Q(u)k(u)du$$

which implies the result. □

Conclusions

In fact, we have

Corollary 3.1.

Suppose that $E[X^{2+\delta}(1)] < \infty$ for some $\delta > 0$. Then (3.1) holds.

Conclusions

Regarding $W(t)$, we first have

Proposition.

Suppose that $B(u_0) < \infty$ for some $u_0 > 1$. Then there exists $\theta_1 > 0$ such that

$$C_1 = \sup_{t \geq 0} E[e^{\theta_1 W(t)}] < \infty.$$

Conclusions

Sketch of proof. By the condition, we have that $F(s,u) < \infty$ for some $u > 1$ and $s > 0$. For convenience, we may assume $K := F(1, u_0) < \infty$ for some $u_0 > 1$. Then for any $t \geq 0$, $F(t+1, u) \leq K$ if $0 \leq u \leq g(t, u_0)$. Further,

$$E[e^{\theta W(t+1)}] \leq K \quad if \quad \theta \leq e^{m(t+1)} \log g(t, u_0).$$

Recall that $g(t, u_0) \downarrow 1$ as $t \uparrow \infty$, by Proposition 2.1,

$$\lim_{t \to \infty} e^{m(t+1)} \log g(t, u_0) = e^m R(u_0) > 0.$$

Therefore, we can choose $\theta_1 > 0$ such that

$$\sup_{t \geq 0} E[e^{\theta_1 W(t)}] \leq K.$$

□

Conclusions

Theorem 3.2.

Suppose that $B(u_0) < \infty$ for some $u_0 > 1$. Then there exists C_2 and $\lambda > 0$ such that

$$P(|W - W(t)| \geq \varepsilon) \leq C_2 e^{-\lambda \varepsilon^{\frac{2}{3}} e^{\frac{mt}{3}}}.$$

Conclusions

Sketch of proof. (i) By Theorem 3.3, we have

$$\phi(\theta) = E[\exp(\theta W)] < \infty, \quad \forall \theta \leq \theta_1.$$

Let $\{W^{(i)};\ i \geq 1\}$ are i.i.d. copies of W, $S_n = \sum_{i=1}^{n}(W^{(i)} - 1)$. We can prove that there exists $\theta_2 > 0\ s.t.$

$$\sup_{|\theta|\leq\theta_2} E[\exp(\frac{\theta S_n}{\sqrt{n}})] \leq e^C$$

where

$$C = \sup_{|u|\leq 1} |\frac{\phi(u)e^{-u} - 1}{u^2}| < \infty.$$

Conclusions

(ii) Note that

$$
\begin{aligned}
W - W(t) &= \lim_{s\to\infty} (W(t+s) - W(t)) \\
&= e^{-mt} \lim_{s\to\infty} \left(e^{-ms} \sum_{j=1}^{X(t)} \xi_{t,j}(s) - X(t) \right) \\
&= e^{-mt} \sum_{j=1}^{X(t)} (W^{(j)} - 1)
\end{aligned}
$$

where $\xi_{t,j}(s)$ is the population size at time $s+t$ of the jth particle among the $X(t)$ particles existing at time t and $W^{(j)}$ is the limit $r.v.$ in the line of descent initiated by jth parent at time t. By the conditional independence,

$$P(W - W(t) > \varepsilon | \sigma(X(s); s \le t)) = \psi(X(t), e^{mt}\varepsilon)$$

where $\psi(k, r) = P(S_k \ge r)$.

Conclusions

(iii) By Markov's inequality,

$$P(S_k \geq r) \leq E\left(e^{\theta_2 \frac{S_k}{\sqrt{k}}}\right) e^{-\theta_2 \frac{r}{\sqrt{k}}} \leq \bar{C} e^{-\theta_2 \frac{r}{\sqrt{k}}}$$

Therefore,

$$P(W - W(t) \geq \varepsilon) = E\psi(X(t), e^{mt}\varepsilon) \leq \bar{C} E\left(e^{-\theta_2 \varepsilon e^{\frac{mt}{2}} \frac{1}{\sqrt{W(t)}}}\right).$$

By Proposition, for $\lambda > 0$,

$$\begin{aligned} E\left[e^{-\frac{\lambda}{\sqrt{W(t)}}}\right] &= \lambda \int_0^\infty e^{-\lambda x} P(e^{\theta_1 W(t)} \geq e^{\frac{\theta_1}{x^2}}) dx \\ &\leq C_1 \int_0^\infty e^{-y} e^{-\frac{\theta_1 \lambda^2}{y^2}} dy. \end{aligned}$$

Thus,

$$P(W - W(t) \geq \varepsilon) \leq \bar{C}C_1 \int_0^\infty e^{-y} e^{-\frac{\theta_1 \lambda_t^2}{y^2}} dy,$$

where $\lambda_t = \theta_2 \varepsilon e^{\frac{mt}{2}}$. However, for $\lambda > 0$,

$$\int_0^\infty e^{-y} e^{-\frac{\lambda^2}{y^2}} dy = \int_0^{\lambda^{2/3}} e^{-y} e^{-\frac{\lambda^2}{y^2}} dy + \int_{\lambda^{2/3}}^\infty e^{-y} e^{-\frac{\lambda^2}{y^2}} dy \leq 2e^{-\lambda^{\frac{2}{3}}}.$$

Hence,

$$P(W - W(t) \geq \varepsilon) \leq 2\bar{C}C_1 e^{-(\sqrt{\theta_1}\theta_2 \epsilon e^{\frac{mt}{2}})^{\frac{2}{3}}} = C_2 e^{-\lambda \varepsilon^{\frac{2}{3}} e^{\frac{mt}{3}}},$$

where $\lambda = (\sqrt{\theta_1}\theta_2)^{\frac{2}{3}}$.
Similar arguments holds for $P(W - W(t) \leq -\varepsilon)$. $\square$

Conclusions

Theorem 3.3.

Suppose that $B(u_0) < \infty$ for some $u_0 > 1$. Then there exists constants C_3 and $\lambda > 0$ such that for all $\varepsilon > 0$, $\alpha > 0$, we can find $0 < I(\varepsilon) < \infty$ such that

$$P\left(\left|\frac{X(t+s)}{X(t)} - e^{ms}\right| > \varepsilon | W \geq \alpha\right) \leq C_3 e^{-\alpha\gamma I(\varepsilon) e^{mt}} + C_2 e^{-\lambda(\alpha(1-\gamma))^{\frac{2}{3}} e^{\frac{mt}{3}}}$$

for $0 < \gamma < 1$. Especially, for the case $\gamma = \frac{1}{2}$,

$$P(|\frac{X(t+s)}{X(t)} - e^{ms}| > \varepsilon | W \geq \alpha) \leq C_4 e^{-\lambda(\frac{\alpha}{2})^{\frac{2}{3}} e^{\frac{mt}{3}}}.$$

Conclusions

Sketch of proof. (i) Note that

$$
\begin{aligned}
& P\left(\left|\frac{X(t+s)}{X(t)}-e^{ms}\right|>\varepsilon|W\geq\alpha\right) \\
&= P\left(\left|\frac{X(t+s)}{X(s)}-e^{ms}\right|>\varepsilon, W\geq\alpha\right)\frac{1}{P(W\geq\alpha)} \\
&= p_\alpha\left[P\left(\left|\frac{X(t+s)}{X(t)}-e^{ms}\right|>\varepsilon, W(t)\leq\alpha\gamma, W\geq\alpha\right)\right] \\
&\quad +p_\alpha\left[P\left(\left|\frac{X(t+s)}{X(t)}-e^{ms}\right|>\varepsilon, W(t)\geq\alpha\gamma, W\geq\alpha\right)\right] \\
&=: p_\alpha(\alpha_{1,t}+\alpha_{2,t}),
\end{aligned}
$$

where $0<\gamma<1, p_\alpha=\frac{1}{P(W\geq\alpha)}$.

Conclusions

(ii) Estimate $\alpha_{1,t}$ and $\alpha_{2,t}$.

By Theorem 3.2,

$$\alpha_{1,t} \leq P(W - W(t) \geq \alpha(1-\gamma)) \leq C_2 e^{-\lambda(\alpha(1-\gamma))^{\frac{2}{3}} e^{\frac{mt}{3}}}.$$

On the other hand, since $E(e^{\theta_1 X(s)}) < \infty$, we can prove that there exist $C_5 > 0$ and $I(\varepsilon) > 0$, $s.t.$

$$\begin{aligned} \alpha_{2,t} &\leq P\left(\left|\frac{X(t+s)}{X(t)} - e^{ms}\right| > \varepsilon, W(t) \geq \alpha\gamma\right) \\ &\leq \sum_{k \geq \alpha\gamma e^{mt}} P(X(t)=k) P\left(\left|\frac{\sum_{i=1}^{k} \xi_{t,i}(s)}{k} - e^{ms}\right| > \varepsilon\right) \\ &\leq C_5 e^{-\alpha\gamma I(\varepsilon) e^{mt}}. \end{aligned}$$

Conclusions

Hence,

$$P\left(\left|\frac{X(t+s)}{X(t)} - e^{ms}\right| > \varepsilon | W \geq \alpha\right)$$
$$\leq \quad p_\alpha \left(C_2 e^{-\lambda(\alpha(1-\gamma))^{\frac{2}{3}} e^{\frac{mt}{3}}} + C_5 e^{-\alpha\gamma I(\varepsilon) e^{mt}}\right).$$

If $\gamma = 1/2$, there exists C_4 and $\lambda > 0$ such that

$$P\left(\left|\frac{X(t+s)}{X(t)} - e^{ms}\right| > \varepsilon | W \geq \alpha\right) \leq C_4 e^{-\lambda(\frac{\alpha}{2})^{\frac{2}{3}} e^{\frac{mt}{3}}},$$

Since the second term $\alpha_{2,t}$ goes to 0 faster than $\alpha_{1,t}$. □

References

ANDERSON,W. (1991).
Continuous-Time Markov Chains: An Applications-Oriented Approach.
Springer-Verlag, New York.

ATHREYA, K.B. (1994).
Large Deviation Rates for Branching Processes–I. Single Type Case.
The Annals of Applied Probability, 4(3):779-790.

ATHREYA,K.B. AND NEY,P.E. (1972).
Branching Processes.
Springer, Berlin.

CHEN,A.Y.,LI,J.P. AND RAMESH, N.I. (2005).
Uniqueness and Extinction of Weighted Markov Branching Processes.
Methodology and Computing in Applied Probability. 7, 489-516.

References

CHOW,Y.S. AND TEICHER H. (1988).
Probability Theory Probability Theory:Independence, Interchangeability, Martingales.
Springer, Newyork.

HARRIS,T.E. (1963).
The theory of branching processes.
Springer, Berlin and Newyork.

KARP,R. AND ZHANG,Y. (1983).
Tail probabilities for finite supercritical braching processes.
Technical Report, Dept. Coputer Science and Engineering, Southern Methodist University, Dallas TX .

MILLER,H.I. AND O'SULLIVAN,J.A. (1992).
Entropies and conbinatorics of random branching processes and context free languages.
IEEE Trans. Inform. Theory, 38, 1292-1311.

References

LI,L.Y. AND LI,J.P. (2018).
Large deviation rates for supercritical branching processes with immigration.
, manuscript.

LIU,J.N. AND ZHANG,M. (2016).
Large deviation for supercritical branching processes with immigration.
Acta Mathematica Sinica, English Series, 32(8):893-900.

SUN,Q. AND ZHANG,M.(2017).
Harmonic moments and large deviations for supercritical branching processes with immigration.
Frontiers of Mathematics in China, 12(5):1201-1220.

Acknowledgements

Thank you!

Email: jpli@mail.csu.edu.cn